A Classical Analog to Entanglement Reversibility

Eric Chitambar,[1] Ben Fortescue,[1] and Min-Hsiu Hsieh[2]

[1] Department of Physics and Astronomy, Southern Illinois University, Carbondale, Illinois 62901, USA
[2] Centre for Quantum Computation & Intelligent Systems (QCIS),
Faculty of Engineering and Information Technology (FEIT),
University of Technology Sydney (UTS), NSW 2007, Australia

(Dated: February 17, 2015)

In this letter we introduce the problem of secrecy reversibility. This asks when two honest parties can distill secret bits from some tripartite distribution p_{XYZ} and transform secret bits back into p_{XYZ} at equal rates using local operation and public communication (LOPC). This is the classical analog to the well-studied problem of reversibly concentrating and diluting entanglement in a quantum state. We identify the structure of distributions possessing reversible secrecy when one of the honest parties holds a binary distribution, and it is possible that all reversible distributions have this form. These distributions are more general than what is obtained by simply constructing a classical analog to the family of quantum states known to have reversible entanglement. An indispensable tool used in our analysis is a conditional form of the Gács-Körner common information.

Resource theories offer a powerful framework for studying what physical processes are possible under a certain class of constraints. For instance, when studying the manipulation of quantum systems, entanglement is identified as a precious resource that cannot be freely generated under local quantum operations and classical communication (LOCC). Inspired by the conceptual successes of entanglement theory, researchers have recently begun applying a resource-theoretic perspective toward the notion of secrecy in classical information theory [1, 2]. For two-party secrecy, one considers tripartite distributions p_{XYZ}: Alice (X) and Bob (Y) share correlations about which, undesirably, Eve (Z) has side information. The distributions are manipulated using local operations and public communication (LOPC), which is the classical analog of LOCC. Just as the ebit $|\Phi\rangle = \sqrt{1/2}(|00\rangle + |11\rangle)$ represents a fundamental unit of entanglement, a secret bit $\Phi_{XY} \cdot q_Z$ represents a fundamental unit of secrecy. Here, $\Phi_{XY}(i,j) = (1/2)\delta_{ij}$ is a perfectly correlated bit while q_Z is an arbitrary and uncorrelated distribution.

Quantum entanglement and classical secrecy share many striking similarities [1–9]. One important similarity lies in the tasks of resource distillation and resource cost. For a bipartite quantum state ρ_{AB}, its *distillable entanglement* $E_D(\rho_{AB})$ quantifies, roughly speaking, the amount of ebits that can be distilled from ρ_{AB} using LOCC [10] (in the many-copy sense), while its *entanglement cost* $E_C(\rho_{AB})$ quantifies the amount of ebits required to generate ρ_{AB} using LOCC [11]. For a distribution p_{XYZ}, its "secrecy content" can analogously be quantified in terms of its distillable key $K_D(p_{XYZ})$ [12, 13] and its key cost $K_C(p_{XYZ})$ [14]. Here, the distillation goal is to obtain secret bits Φ_{XY} from p_{XYZ}, while the formation goal is simulate p_{XYZ} using Φ_{XY} and public communication. Compared to entanglement theory, much less is known about the relationship between K_D and K_C, except for the expected hierarchy $K_C \geq K_D$ [14].

With the inequality $K_C \geq K_D$, classical secrecy can be given a thermodynamic interpretation similar to entanglement [15, 16]. By the second law of thermodynamics, a heat engine cannot do more work when transferring heat from one temperature bath to a lower one than the work required to perform the reverse refrigeration process. Likewise, $K_C(p_{XYZ}) \geq K_D(p_{XYZ})$ means that an LOPC protocol is not able to distill more secret bits from p_{XYZ} than the secret bits needed to perform the reverse formation process. Any distribution for which this inequality is tight can thus be regarded as the secrecy analog of a reversible heat engine. The *secrecy reversibility problem* asks what distributions satisfy $K_C(p_{XYZ}) = K_D(p_{XYZ})$.

To begin tackling this problem, it is instructive to first consider the quantum scenario. It is well-known that all bipartite quantum pure states demonstrate entanglement reversibility: any pure state can be concentrated into an EPR state $|\Phi\rangle$ and diluted back to the original state at equal rates [17]. Thus, a natural starting place to find reversible secrecy is with a classical analog to quantum pure states. Collins and Popescu have investigated [1] one such analog based on an embedding of $p(x,y,z)$ into a tripartite quantum state given by

$$|\Psi\rangle_{ABE} = \sum_{x,y,z} \sqrt{p(x,y,z)}|xyz\rangle. \tag{1}$$

If Alice and Bob's reduced state in $|\Psi\rangle$ is pure, then $|\Psi\rangle$ can always be expressed as $|\Psi\rangle = \sum_{j,z} \sqrt{p(j)q(z)}|\alpha_j\beta_j\rangle|z\rangle$, where $|\alpha_j\rangle$ and $|\beta_j\rangle$ are Schmidt basis vectors. With this motivation, Collins and Popescu have proposed distributions of the form $p(x,y,z) = \delta_{xy}p(x)q(z)$ to be the classical analog to quantum pure states (another type of analog has also been proposed in the literature [18]). Actually, we can generalize the Collins-Popescu class of distributions to include distributions of the form

$$p(x,y,z)=\sum_j p(x|j)p(y|j)p(j)q(z), \tag{2}$$

where $p(x|j)p(x|j')=p(y|j)p(y|j')=0$ if $j\neq j'$. A quantum embedding of any such distribution *à la* Eq. (1) recovers a pure state for Alice and Bob with Schmidt basis vectors $|\alpha_j\rangle=\sum_x\sqrt{p(x|j)}|x\rangle$ and $|\beta_j\rangle=\sum_y\sqrt{p(y|j)}|y\rangle$. We refer to any distribution having the the form of Eq. (2) as *secret block independent* (SBI), and they may also be considered as a type of "classical pure state." Like the Collins-Popescu distributions, the theory of single-copy state transformations can be constructed for SBI states analogous to pure quantum states [1]. Furthermore, just as pure quantum states possess reversible entanglement, SBI distributions possess reversible secrecy, as will be shown below (also see footnote 12 in [7]).

However, it turns out that a much richer class of distributions beyond SBI also demonstrate secrecy reversibility. To begin studying this class, we first recall a well-known upper bound on $K_D(p_{XYZ})$ referred to as the *intrinsic information* of p_{XYZ} [19]. This quantity is given by

$$I(X:Y\downarrow Z):=\min I(X:Y|\overline{Z}), \tag{3}$$

where the minimization is taken over over all auxiliary variables $\overline{Z}$ such that $XY-Z-\overline{Z}$ forms a Markov chain [20]. Using the definition of key cost, Renner and Wolf were able to prove that $K_C(p_{XYZ})\geq I(X:Y\downarrow Z)$ [14], and thus

$$K_D(p_{XYZ})\leq I(X:Y\downarrow Z)\leq K_C(p_{XYZ}). \tag{4}$$

Consequently, we can split the secrecy reversibility problem into two separate questions: (1) when does $K_C(p_{XYZ})=I(X:Y\downarrow Z)$, and (2) when does $K_D(p_{XYZ})=I(X:Y\downarrow Z)$? We answer the first question below and reference certain results from Ref. [21] where we have recently studied the second question. However, before doing so, we introduce a variety of distribution classes based on the notion of a conditional common function since these classes will play a central role in our analysis of reversible secrecy.

Common Functions and UBI-PD↓ Distributions. For distribution p_{XY}, a *maximal common function* is a variable J_{XY} such that

$$H(J_{XY})=\max_K\{H(K):0=H(K|X)=H(K|Y)\}. \tag{5}$$

The value $H(J_{XY})$ has been identified by Gács and Körner as the *common information* between X and Y [22]. It can be shown that for every p_{XY}, the variable J_{XY} is unique up to a relabeling of its range (see Supplemental Material). Note that an SBI distribution can be equivalently characterized by the entropic condition $I(X:Y|J_{XY})=0$, and $H(J_{XY})=I(X:Y)$ for these distributions [22].

For a tripartite distribution p_{XYZ}, we will denote a maximal common function of the conditional distribution $p_{XY|Z=z}$ by $J_{XY|Z=z}$. Then, a *maximal conditional common function* $J_{XY|Z}$ is just a collection of maximal common functions $\{J_{XY|Z=z}:p(z)>0\}$. Again, the variable $J_{XY|Z}$ is unique up to relabeling. We say that a distribution p_{XYZ} is *block independent* (BI) if $I(X:Y|J_{XY|Z}Z)=0$; equivalently, if the distribution decomposes as

$$p(x,y,z)=\sum_{z\in\mathcal{Z}}\sum_{J_{XY|Z=z}=j}p(x|z,j)p(y|z,j)p(j,z), \tag{6}$$

where $p(x|z,j)p(x|z,j')=0$ and $p(y|z,j)p(y|z,j')=0$ for $j\neq j'$. Obviously SBI distributions are simply BI with an uncorrelated Eve. A distribution is said to be *uniform block independent* (UBI) if it is block independent, and there exist local coarse-graining maps $K_X(X)$ and $K_Y(Y)$ such that $Pr[J_{XY|Z}=K_X=K_Y]=1$ for some maximal common function $J_{XY|Z}$. In other words, Alice and Bob can determine the value for $J_{XY|Z}$ simply by consulting their local variable. With many copies of a UBI distribution, secret key can be distilled via privacy amplification at an optimal rate $H(J_{XY|Z}|Z)=I(X:Y|Z)$ [12, 23].

However, in general $J_{XY|Z}$ will be unknown to Alice and Bob unless they engage in public communication. A public communication protocol is a sequence of public messages $M=(M_1,M_2,\cdots,M_r)$ such that M_k is a function of both $M_{k-1}\cdots M_1$ and X (resp. Y) when k is odd (resp. even). At the end of these exchanges, the new object of interest becomes $J_{XY|ZM}$, which is a maximal conditional common function for the distribution $p_{(XM)(YM)(ZM)}$. It can easily be proven that when p_{XYZ} is BI, so is $p_{(MX)(MY)(ZM)}$, and furthermore

$$I(X:Y|ZM)=I(X:Y|Z)-I(M:J_{XY|Z}|Z) \tag{7}$$

(see Supplemental Material). This equation formalizes the intuitive idea that messages M will decrease Alice and Bob's average conditional common information unless, from Eve's perspective, the messages are independent of $J_{XY|Z}$.

With this motivation, we say p_{XYZ} is *uniform block independent under public discussion* (UBI-PD) if it is BI and there is a public communication protocol generating messages M such that $p_{(MX)(MY)(ZM)}$ is UBI and $I(M:J_{XY|Z}|Z)=0$. Thus, UBI-PD distributions have a distillation rate of $H(J_{XY|ZM}|ZM)$, which by Eq. (7) is equal to $H(J_{XY|Z}|Z)=I(X:Y|Z)$. We say a distribution belongs to the class UBI-PD↓ if there exists a channel $\overline{Z}|Z$ such that $p_{XY|\overline{Z}}$ is UBI with the required public communication M also satisfying $I(Z:J_{XY|\overline{Z}}|M\overline{Z})=0$. This latter condition assures

$X \longrightarrow$

$Y \downarrow$

$Z{=}0$	0	1	2
0	1/2	·	·
1	·	1/8	·
2	·	·	3/8

$\overline{Z}|Z$

$\overline{Z}{=}0$	0	1	2
0	1/2	·	·
1	·	1/8	1/8
2	·	1/8	1/8

$Z{=}1$	0	1	2
0	1/2	·	·
1	·	1/8	5/32
2	·	5/32	1/16

← i.e. $p_{XY|Z}(2,2|1) = \frac{1}{16}$

FIG. 1. A UBI-PD↓ distribution where $p_Z(0) = 1/5$, $p_Z(1) = 4/5$ and $\overline{Z}$ is a full coarse-graining of Z. In this simplified example, no communication is needed for Alice and Bob to both generate $J_{XY|\overline{Z}}$. Note that p_{XYZ} itself is not BI.

that $K_D(p_{XYZ}) = H(J_{XY|\overline{Z}}|\overline{Z}) = I(X : Y \downarrow Z)$. Indeed, for every UBI-PB↓ distribution, $J_{XY|\overline{Z}}$ becomes shared randomness under communication M. Thus, an achievable key rate is

$$H(J_{XY|\overline{Z}}|ZM) = H(J_{XY|\overline{Z}}|\overline{Z}M) = H(J_{XY|\overline{Z}}|\overline{Z}),$$

where the first equality follows from $I(Z : J_{XY|\overline{Z}}|M\overline{Z}) = 0$ and the second from $I(M : J_{XY|\overline{Z}}|\overline{Z}) = 0$. Fig. 1 depicts a UBI-PD↓ distribution. We encourage the reader to visit the Supplemental Material for a comparative picture of the various distribution classes identified here.

When does $K_C(p_{XYZ}) = I(X : Y \downarrow Z)$? This question can be answered using the formula for key cost as computed by Winter, a significant result on its own since no single-letter expression is known for $K_D(p_{XYZ})$.

Lemma 1 (Winter [24]). *For a distribution p_{XYZ},*

$$K_C(p_{XYZ}) = \min I(XY : W|\overline{Z}), \tag{8}$$

where the minimization is over all auxiliary variables W and $\overline{Z}$ which satisfy $XY - Z - \overline{Z}$ and $X - W\overline{Z} - Y$.

Our task will now be to reproduce Renner and Wolf's result that $K_C(p_{XYZ}) \geq I(X : Y \downarrow Z)$ directly from Winter's formula (8). In doing so, we will obtain a structure condition for when $K_C(p_{XYZ}) = I(X : Y \downarrow Z)$.

Lemma 2. *For the distribution p_{XYZ}, $K_C(p_{XYZ}) \geq I(X : Y \downarrow Z)$. Equality is obtained iff $p_{XY\overline{Z}}$ is BI, where $\overline{Z}|Z$ is the minimizer in $I(X : Y \downarrow Z)$. When equality holds, $K_C(p_{XYZ}) = I(X : Y \downarrow Z) = H(J_{XY|\overline{Z}}|\overline{Z})$.*

Proof. Let $XYZW\overline{Z}$ satisfy the minimization in Eq. (8). Then we have the following chain of inequalities:

$$\begin{aligned} K_C(p_{XYZ}) = I(XY : W|\overline{Z}) &\geq I(X : W|\overline{Z}) \\ &\geq I(X : Y|\overline{Z}) \geq I(X : Y \downarrow Z). \end{aligned} \tag{9}$$

The first inequality follows from the fact that $I(Y : W|X\overline{Z}) \geq 0$ with equality obtained iff $W - X\overline{Z} - Y$; the second inequality is the data-processing inequality applied to $X - W\overline{Z} - Y$ with equality obtained iff $X - Y\overline{Z} - W$; and the third inequality follows from the definition of intrinsic information.

For the equality conditions, consider when $X - Y\overline{Z} - W$ and $Y - X\overline{Z} - W$. This so-called "conditional double Markov chain" can only be satisfied if $I(XY : W|J_{XY|\overline{Z}}\overline{Z}) = 0$ (see Supplemental Material). Using this we upper bound the key cost by

$$\begin{aligned} I(XY : W|\overline{Z}) &= I(XYJ_{XY|\overline{Z}} : W|\overline{Z}) \\ &= I(J_{XY|\overline{Z}} : W|\overline{Z}) \leq H(J_{XY|\overline{Z}}|\overline{Z}). \end{aligned} \tag{10}$$

Since $J_{XY|Z}$ is both a function of X and Y given Z, it is easy to show $H(J_{XY|\overline{Z}}|\overline{Z}) \leq I(X : Y|\overline{Z})$, with equality iff $I(X : Y|\overline{Z}J_{XY|\overline{Z}}) = 0$. Hence demanding that $I(XY : W|\overline{Z}) = I(X : Y|\overline{Z})$ gives the necessary conditions $H(J_{XY|\overline{Z}}|W\overline{Z}) = 0$ and $I(X;Y|J_{XY|\overline{Z}}\overline{Z}) = 0$.

Conversely, if $p_{XY\overline{Z}}$ is block independent and $I(X : Y \downarrow Z) = I(X : Y|\overline{Z})$, then choose $\overline{Z}$ and $W = J_{XY|\overline{Z}}$ in the minimization of Eq. (8) to obtain $K_C(p_{XYZ}) = I(X : Y \downarrow Z)$. □

A Class of Reversible Distributions. We have seen that $K_D(p_{XYZ}) = I(X : Y \downarrow Z)$ for UBI-PD↓ distributions. Since these distributions admit a channel $\overline{Z}|Z$ with $p_{XY\overline{Z}}$ being BI, Lemma 2 gives that $K_D(p_{XYZ}) = K_C(p_{XYZ})$ for every UBI-PD↓ distribution. We have thus identified a family of distributions possessing reversible secrecy, and we conjecture that this family completely characterizes secrecy reversibility in the classical setting. The conjecture obviously holds true for any distribution with $0 = K_C(p_{XYZ}) = K_D(p_{XYZ})$ since $K_C(p_{XYZ}) = 0$ implies $I(X : Y \downarrow Z) = 0$ by Lemma 2, and any distribution satisfying the latter condition is UBI-PB↓ by definition. The conjecture can also be shown as true for distributions satisfying $\min\{|\mathcal{X}|, |\mathcal{Y}|\} = 2$.

Theorem 1. *If $\min\{|\mathcal{X}|, |\mathcal{Y}|\} = 2$, then $K_C(p_{XYZ}) = K_D(p_{XYZ})$ iff p_{XYZ} is UBI-PD↓.*

Proof. Here we prove the theorem for when $|\mathcal{X}| = |\mathcal{Y}| = 2$, and the more general case is handled in the Supplemental Material. Crucial to our argument is a necessary structural condition recently proven for distributions satisfying $K_D(p_{XYZ}) = I(X : Y|Z)$ [21].

Proposition 1 ([21]). *When $|X| = |Y| = 2$ and there exists a pair (x, y) such that $p(x, y|z_1)p(x|z_0)p(y|z_0) > 0$ but $p(x, y|z_0) = 0$ for some $z_0, z_1 \in \mathcal{Z}$, then $K_D(p_{XYZ}) < I(X : Y|Z)$.*

Continuing with the proof of Theorem 1 in the 2×2 case, from the previous discussion it suffices to prove necessity when $K_C(p_{XYZ}) = K_D(p_{XYZ}) > 0$. Then by Lemma

2, for some $\overline{Z}|Z$, $p_{XY\overline{Z}}$ must be block independent and $K_D(p_{XYZ}) = I(X:Y|\overline{Z})$. However, since $K_D(p_{XYZ}) \leq K_D(p_{XY\overline{Z}}) \leq I(X:Y|\overline{Z})$, we see that $K_D(p_{XY\overline{Z}}) = I(X:Y|\overline{Z})$. Then from Proposition 1, the structure of BI distributions, and the fact that $H(J_{XY|\overline{Z}}|\overline{Z}=z) > 0$ for some z, we have that $H(X|Y) = H(Y|X) = 0$; i.e. $p_{XY\overline{Z}}$ is UBI and, up to a relabeling, has the form $p(x,y,z) = \delta_{xy}[xq(z) + (1-x)(1-q(z))]$. Since $\overline{Z}$ is obtained by processing Z, $p_{XY\overline{Z}}$ can have this correlated form only if p_{XYZ} likewise does. Thus, p_{XYZ} is UBI. □

Reversible Distributions Embedded in Quantum States. We now consider embedding reversible distributions into quantum states as in Eq. (1). In particular, we focus on distributions with $|\mathcal{X}| = |\mathcal{Y}| = 2$ so that the corresponding $\rho_{AB} := \mathrm{Tr}_E |\Psi\rangle\langle\Psi|_{ABE}$ is a two-qubit state. We can make a comparison between the secret key of the underlying distribution and the entanglement of the embedded quantum state using an analytic formula for the entanglement of formation E_F[25]. The following relatively straightforward calculation is carried out in the Supplemental Material.

Theorem 2. *For reversible p_{XYZ} with $|\mathcal{X}| = |\mathcal{Y}| = 2$ and $K_D(p_{XYZ}) > 0$:*

$$\begin{aligned} K_D(p_{XYZ}) &= \sum_{z\in\mathcal{Z}} p(z)\mathsf{E}\left(2\sqrt{p(0|z)p(1|z)}\right) \\ E_F(\rho_{AB}) &= \mathsf{E}\left(2\sum_{z\in\mathcal{Z}} p(z)\sqrt{p(0|z)p(1|z)}\right), \end{aligned} \tag{11}$$

where $\mathsf{E}(x) := h(\frac{1}{2}[1-\sqrt{1-x^2}])$ is strictly convex in x for $h(x) := -x\log x - (1-x)\log(1-x)$. The equality $K_D(p_{XYZ}) = E_F(\rho_{AB})$ holds iff $H(X|Z=z)$ is constant for all $z \in \mathcal{Z}$.

It is natural to wonder whether a quantum state with an embedded reversible distribution will likewise possess reversible entanglement. However, one can already see in two qubits that this will not be true in general. Every two-qubit embedded ρ_{AB} with $K_D(p_{XYZ}) > 0$ will take the form $\rho_{AB} = \sum_z \sum_{j,j'=0}^{1} p(z)\sqrt{p(j|z)p(j'|z)}|jj\rangle\langle j'j'|$. This is a so-called maximally-correlated state for which entanglement reversibility is known to be lacking whenever ρ_{AB} is not pure [26, 27]. In fact, $E_F(\rho_{AB})$ is additive for the states of Theorem 2 [28]. Thus,

Corollary 1. *When $|\mathcal{X}| = |\mathcal{Y}| = 2$, any distribution with nonzero reversible secrecy will have nonzero reversible entanglement when embedded in a quantum state iff the embedded state is pure.*

Returning to Reversible Entanglement. We motivated our investigation into reversible secrecy by considering reversible entanglement in quantum pure states and asking for a classical analog. This led to the proposal of SBI distributions as being a type of "classical pure state." Beyond pure states, the only known quantum mixed states demonstrating entanglement reversibility are the so-called locally-flagged states [26, 27, 29, 30]. By generalizing the type of states presented in [29], we say that σ_{AB} is an *LOCC-flagged* state if there exists an LOCC instrument $(\mathcal{L}_m)_m$ (i.e. a collection of CP maps generated by an LOCC protocol [31]), with m enumerating the different possible public messages of the protocol, such that (i) $\sigma = \sum_m \mathcal{L}_m(\sigma)$ and (ii) $\frac{1}{p(m)}\mathcal{L}_m(\sigma) = |\varphi_m\rangle\langle\varphi_m|$ is pure, where $p(m) = \|\mathcal{L}_m(\sigma)\|_1$. For such states, $E_C(\sigma) = E_D(\sigma) = \sum_m p(m)S(\mathrm{Tr}_A |\varphi_m\rangle\langle\varphi_m|)$.

What is the classical analog of LOCC-flagged mixed states? Care must be taken since in the definition of key cost, Eve must be able to use her part of p_{XYZ} to simulate whatever public communication Alice and Bob use to generate their parts of p_{XYZ} in a formation protocol [14]. Given the identification of an SBI distribution as a classical pure state, we say distribution p_{XYZ} is an *LOPC-flagged* state if there exists an LOPC instrument $(\mathcal{L}_m)_m$ (i.e. a collection of substochastic maps generated by an LOPC protocol), with m enumerating the different public messages of the protocol, such that (i) $p_{XYZ} = \sum_m \mathcal{L}_m(p_{XYZ})$, (ii) $\frac{1}{p(m)}\mathcal{L}_m(p_{XYZ}) = p(x,y|m)p(z|m)$ is SBI, where $p(m) = \|\mathcal{L}_m(p_{XYZ})\|_1$, and (iii) $p(z|m)p(z|m') = 0$ for $m \neq m'$. This is formally analogous to the quantum scenario except for condition (iii), which captures the ability for Eve to reproduce the public communication from her information Z. Any LOPC-flagged classical state takes the form

$$p(x,y,z) = \sum_{M=m} p(x,y|m)p(z|m)p(m) \tag{12}$$

where M is generated by a public communication protocol with $I(X:Y|J_{XY|M}, M) = 0$ and $H(M|Z) = 0$. It immediately follows from definition that these distributions are UBI-PD, but the converse is not true.

Conclusions. We have presented a class of distributions UBI-PD↓ that are conjectured to fully characterize reversible secrecy. Despite the complexity of these distributions, validity of this conjecture would mean that reversibility of some distribution could be decided by a single-copy analysis. Turning back to the analogous problem of entanglement reversibility in quantum states, one might then likewise hope for a solution on the single-copy level. Only LOCC-flagged mixed states are known to possess entanglement reversibility, and these can indeed be identified by having a particular single-copy structure. We have proposed a classical analog to LOCC-flagged states that likewise possess reversible secrecy, but these do not constitute the full set of reversible states. Therefore, if only LOCC-flagged quantum states possess entanglement reversibility, then the analogous statement for secrecy in classical states would not be true. On the other hand, if entanglement and secrecy are truly on equal footing in terms of reversibility characters, then our findings might suggest the existence of reversible entanglement beyond LOCC-flagged states.

EC thanks Matthias Christandl for a helpful discussion on the intrinsic information and key cost. EC was supported by the National Science Foundation (NSF) Early CAREER Award No. 1352326. MH is supported by an ARC Future Fellowship under Grant FT140100574.

[1] D. Collins and S. Popescu, Phys. Rev. A **65**, 032321 (2002).

[2] K. Horodecki, M. Horodecki, P. Horodecki, and J. Oppenheim, Foundations of Physics **35**, 2027 (2005).

[3] N. Gisin, R. Renner, and S. Wolf, Algorithmica **34**, 389 (2002).

[4] A. Acín, L. Masanes, and N. Gisin, Phys. Rev. Lett. **91**, 167901 (2003).

[5] A. Acín and N. Gisin, Phys. Rev. Lett. **94**, 020501 (2005).

[6] M. Christandl, A. Ekert, M. Horodecki, P. Horodecki, J. Oppenheim, and R. Renner, in *Theory of Cryptography*, Lecture Notes in Computer Science, Vol. 4392, edited by S. Vadhan (Springer Berlin Heidelberg, 2007) pp. 456–478.

[7] J. Oppenheim, R. W. Spekkens, and A. Winter, "A classical analogue of negative information," (2008), accepted into Phys. Rev. Lett., arXiv:quant-ph/0511247v2.

[8] J. Bae, T. Cubitt, and A. Acín, Phys. Rev. A **79**, 032304 (2009).

[9] M. Ozols, G. Smith, and J. A. Smolin, Phys. Rev. Lett. **112**, 110502 (2014).

[10] E. M. Rains, Phys. Rev. A **60**, 173 (1999).

[11] P. M. Hayden, M. Horodecki, and B. M. Terhal, J. Phys. A: Math. Gen. **34**, 6891 (2001).

[12] R. Ahlswede and I. Csiszár, Information Theory, IEEE Transactions on **39**, 1121 (1993).

[13] U. Maurer, Information Theory, IEEE Transactions on **39**, 733 (1993).

[14] R. Renner and S. Wolf, in *Advances in Cryptology EUROCRYPT 2003*, Lecture Notes in Computer Science, Vol. 2656 (Springer Berlin Heidelberg, 2003) pp. 562–577.

[15] S. Popescu and D. Rohrlich, Phys. Rev. A **56**, R3319 (1997).

[16] M. Horodecki, J. Oppenheim, and R. Horodecki, Phys. Rev. Lett. **89**, 240403 (2002).

[17] C. H. Bennett, H. Bernstein, S. Popescu, and B. Schumacher, Phys. Rev. A **53**, 2046 (1996), quant-ph/9511030.

[18] In Ref. [7], the authors introduce another class of distributions, called *bi-disjoint*, which they propose as a different analog to quantum pure states. In terms of the notation used here, a distribution is bi-disjoint iff $I(XY:Z|J_{(XY)Z})=0$, where $J_{(XY)Z}$ is the common information between Alice-Bob (jointly) and Eve. It is shown in [7] that bi-disjoint distributions behave like quantum pure states for the task of state merging. However, in general they fail to possess reversible secrecy, nor do they behave like quantum pure states for single-copy state transformations. It therefore seems that classical analogies to quantum pure states can only be drawn *with respect to specific information-theoretic tasks*, a conclusion already implicitly acknowledged in [7]. For the task of resource reversibility, SBI distributions are the more appropriate analog to quantum pure states.

[19] U. Maurer and S. Wolf, Information Theory, IEEE Transactions on **45**, 499 (1999).

[20] Recall, a triple of random variables ABC ranging over $\mathcal{A}\times\mathcal{B}\times\mathcal{C}$ form a Markov chain $A{-}B{-}C$ if $p(a|bc)=p(a|b)$ for all $(a,b,c)\in\mathcal{A}\times\mathcal{B}\times\mathcal{C}$ such that $p(b,c)>0$. In entropic terms, this is equivalent to the vanishing of the conditional mutual information: $I(A,C|B)=0$.

[21] E. Chitambar, B. Fortescue, and M.-H. Hsieh, "Distributions attaining secret key at a rate of the conditional mutual information," (2014), manuscript in preparation.

[22] P. Gács and J. Körner, Problems of Control and Information Theory **2**, 149 (1973).

[23] C. Bennett, G. Brassard, C. Crepeau, and U. Maurer, Information Theory, IEEE Transactions on **41**, 1915 (1995).

[24] A. Winter, in *Information Theory, 2005. ISIT 2005. Proceedings. International Symposium on* (2005) pp. 2270–2274.

[25] W. K. Wootters, Phys. Rev. Lett. **80**, 2245 (1998).

[26] M. F. Cornelio, M. C. de Oliveira, and F. F. Fanchini, Phys. Rev. Lett. **107**, 020502 (2011).

[27] K. G. H. Vollbrecht, R. F. Werner, and M. M. Wolf, Phys. Rev. A **69**, 062304 (2004).

[28] G. Vidal, W. Dür, and J. I. Cirac, Phys. Rev. Lett. **89**, 027901 (2002).

[29] P. Horodecki, R. Horodecki, and M. Horodecki, Acta Physica Slovaca **48**, 141 (1998).

[30] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Rev. Mod. Phys. **81**, 865 (2009).

[31] E. Chitambar, D. Leung, L. Maninska, M. Ozols, and A. Winter, Communications in Mathematical Physics **328**, 303 (2014).

[32] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems* (Cambridge University Press, Cambridge, UK, 2011).

Properties of the Gács-Körner Common Information

In this appendix, we prove the variable J_{XY} that maximizes Eq. (5) is unique up to relabeling of its range. To do this we give an alternative characterization of J_{XY}, directly reminiscent of that in [22]. Let X and Y be random variables over finite sets $\mathcal{X}$ and $\mathcal{Y}$ respectively, with joint distribution p_{XY}. A *common partitioning of length* t for XY are pairs of subsets $(\mathcal{X}_i,\mathcal{Y}_i)_{i=1}^t$ such that

(i) $\mathcal{X}_i\cap\mathcal{X}_j=\mathcal{Y}_i\cap\mathcal{Y}_j=\emptyset$ for $i\neq j$,

(ii) $p(\mathcal{X}_i|\mathcal{Y}_j)=p(\mathcal{Y}_i|\mathcal{X}_j)=\delta_{ij}$, and

(iii) if $(x,y)\in\mathcal{X}_i\times\mathcal{Y}_i$ for some i, then $p_X(x)p_Y(y)>0$.

For a given common partitioning, we refer to the subsets $\mathcal{X}_i\times\mathcal{Y}_i$ as the "blocks" of the partitioning. The subscript i merely serves to label the different blocks, and for any fixed labeling, we associate a random variable $J(X,Y)$ such that $J(x,y)=i$ if $(x,y)\in\mathcal{X}_i\times\mathcal{Y}_i$. Note that each party can determine the value of J from their local information, and it is therefore called a *common function* of X and Y [22]. A *maximal common partitioning* is a common partitioning of greatest length.

Proposition 2.

(a) Every pair of finite random variables XY has a unique maximal common partitioning.

(b) Variable J_{XY} satisfies

$$H(J_{XY})=\max_K\{H(K):0=H(K|X)=H(K|Y)\}$$

iff J_{XY} is a common function for the maximal common partitioning of XY.

Proof. (a) Trivially $\mathcal{X}\times\mathcal{Y}$ gives a common partitioning of length one, and any common partitioning cannot have length exceeding $\min\{|\mathcal{X}|,|\mathcal{Y}|\}$; hence a maximal common partitioning exists. To prove uniqueness, suppose that $(\mathcal{X}_i,\mathcal{Y}_i)_{i=1}^t$ and $(\mathcal{X}'_i,\mathcal{Y}'_i)_{i=1}^t$ are two maximal common partitionings. If they are not equivalent, then there must exist some subset, say $\mathcal{X}_{i_0}$ such that $\mathcal{X}_{i_0}\subset\cup_{\lambda=1}^K\mathcal{X}'_\lambda$ in which $\mathcal{X}_{i_0}\cap\mathcal{X}'_\lambda\neq\emptyset$ for $\lambda=1,\cdots,K\geq 2$. Choose any such $\mathcal{X}'_{\lambda_0}$ from this collection and define the new sets $R_{i_0}=\mathcal{X}_{i_0}\cap\mathcal{X}'_{\lambda_0}$ and $\tilde{R}_{i_0}=\mathcal{X}_{i_0}\setminus\mathcal{X}'_{\lambda_0}$, which are both nonempty since $k\geq 2$ and the $\mathcal{X}_\lambda$ are disjoint. However, we also have the properties

$$x\in\mathcal{X}_{i_0}\Rightarrow p(\mathcal{Y}_{i_0}|x)=1; \qquad x\in\mathcal{X}'_{\lambda_0}\Rightarrow p(\mathcal{Y}'_{\lambda_0}|x)=1;$$
$$x\notin\mathcal{X}_{i_0}\Rightarrow p(\mathcal{Y}_{i_0}|x)=0; \qquad x\notin\mathcal{X}'_{\lambda_0}\Rightarrow p(\mathcal{Y}'_{\lambda_0}|x)=0.$$

(Here we are implicitly using condition (iii) in the above definition by assuming that $p(x)>0$ thereby defining conditional distributions). Therefore, $p(S_{i_0}|R_{i_0})=p(\tilde{S}_{i_0}|\tilde{R}_{i_0})=1$ and $p(S_{i_0}|\tilde{R}_{i_0})=p(\tilde{S}_{i_0}|R_{i_0})=0$, where $S_{i_0}=\mathcal{Y}_{i_0}\cap\mathcal{Y}'_{\lambda_0}$ and $\tilde{S}_{i_0}=\mathcal{Y}_{i_0}\setminus\mathcal{Y}'_{\lambda_0}$. A similar argument shows that $p(R_{i_0}|S_{i_0})=p(\tilde{R}_{i_0}|\tilde{S}_{i_0})=1$ and $p(R_{i_0}|\tilde{S}_{i_0})=p(\tilde{R}_{i_0}|S_{i_0})=0$. Hence, $(\mathcal{X}_i,\mathcal{Y}_i)_{i\neq i_0}^t\bigcup(S_{i_0},R_{i_0})\bigcup(\tilde{S}_{i_0},\tilde{R}_{i_0})$ is a common partitioning of length $t+1$. But this is a contradiction since $(\mathcal{X}_i,\mathcal{Y}_i)_{i=1}^t$ is a maximal common decomposition.

(b) Suppose that K satisfies $0=H(K|X)=H(K|Y)$ so that $K=f(X)=g(Y)$ for some functions f and g. It is clear that f and g must be constant-valued for any pair of values taken from same block $\mathcal{X}_i\times\mathcal{Y}_i$ in the maximal common partitioning of XY. Hence the maximum possible entropy of K is then attained iff f and g take on a different value for each block in this partitioning. □

We now turn to the conditional common information $J_{XY|Z}$. We are specifically interested in the how this information evolves under LOPC for block independent distributions. The following provides a derviation of Eq. (7).

Proposition 3. *If p_{XYZ} is BI, then so is $p_{(MX)(MY)(ZM)}$. Moreover,*

$$I(X:Y|ZM)=I(X:Y|Z)-I(M:J_{XY|Z}|Z). \tag{13}$$

Proof. For a general distribution p_{XYZ}, it is easy to see that $H(K|Z)\leq I(X:Y|Z)$ whenever $H(K|XZ)=H(K|YZ)=0$. Equality is obtained iff p_{XYZ} is BI, and by uniqueness of the maximal conditional common function, we have that $K=J_{XY|Z}$ up to relabeling.

Now, suppose that p_{XYZ} is BI and Alice locally generates message M_1 so that $YZ-X-M_1$. Then

$$\begin{aligned}I(X:Y|ZM_1)&=I(M_1X:Y|Z)-I(M_1:Y|Z)\\&=I(X:Y|Z)-I(M_1:J_{XY|Z}Y|Z)\\&=I(X:Y|Z)-I(M_1:J_{XY|Z}|Z)\\&=H(J_{XY|Z}|Z)-[H(J_{XY|Z}|Z)-H(J_{XY|Z}|ZM_1)]\\&=H(J_{XY|Z}|ZM_1).\end{aligned}\tag{14}$$

Since $H(J_{XY|Z}|XM_1)=H(J_{XY|Z}|YM_1)=0$, by the above discussion it follows that $p_{(XM)(YM)(ZM)}$ is BI and $J_{XY|ZM_1}$ is essentially equivalent to $J_{XY|Z}$; i.e. up to relabeling $J_{XY|Z=z}=J_{XY|Z=z,M_1=m}$ for all m. The third line of Eq. (14) gives us the desired equality in the proposition for message M_1. Proceeding by induction proves the full statement for a full message M generated by an arbitrarily long communication protocol.

□

A Hierarchy of Distribution Classes

We review the various distributions classes introduced in the paper and give different examples. The hierarchy of the distributions is the following:

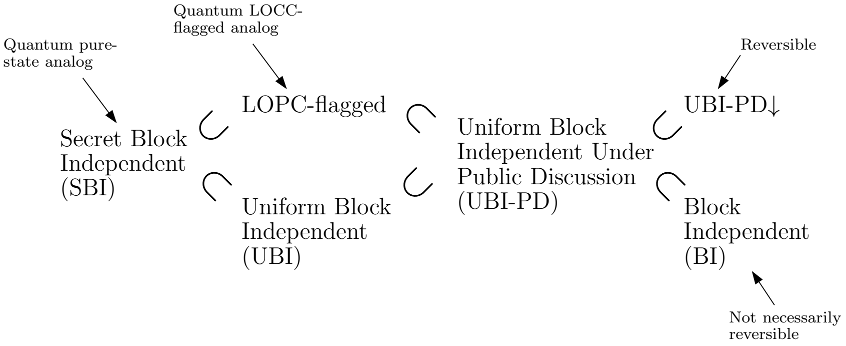
Quantum pure-state analog
Quantum LOCC-flagged analog
Reversible
Secret Block Independent (SBI)
LOPC-flagged
Uniform Block Independent (UBI)
Uniform Block Independent Under Public Discussion (UBI-PD)
UBI-PD↓
Block Independent (BI)
Not necessarily reversible

FIG. 2. A hierarchy of distribution classes and their relation to classes of reversible quantum states.

Example Distributions:

(a) LOPC-flagged:

$X \longrightarrow$, $Y \downarrow$

$Z=0$	0	1	2
0	1/8	1/8	·
1	1/8	1/8	·
2	·	·	1/2

$p_{XY|Z=0}$

$Z=1$	0	1	2
0	1/8	1/8	·
1	1/8	1/8	·
2	·	·	1/2

$p_{XY|Z=1}$

$Z=2$	3	4	5
0	1/4	·	·
1	·	1/4	·
2	·	·	1/2

$p_{XY|Z=2}$

(b) Uniform Block Independent (UBI):

$X \longrightarrow$, $Y \downarrow$

$Z=0$	0	1	2
0	1/8	1/8	·
1	1/8	1/8	·
2	·	·	1/2

$p_{XY|Z=0}$

$Z=1$	0	1	2
0	1/6	1/6	·
1	1/6	1/6	·
2	·	·	1/3

$p_{XY|Z=1}$

$Z=2$	3	4	5
0	1/4	1/4	·
1	·	·	·
2	·	·	1/2

$p_{XY|Z=2}$

(c) Uniform Block Independent Under Public Discussion (UBI-PD):

$X \longrightarrow$, $Y \downarrow$

$Z=0$	0	1	2
0	1/8	1/8	·
1	1/8	1/8	·
2	·	·	1/2

$p_{XY|Z=0}$

$Z=1$	0	1	2
0	1/6	1/6	·
1	1/6	1/6	·
2	·	·	1/3

$p_{XY|Z=1}$

$Z=2$	3	4	5
0	1/2	·	·
1	·	1/8	1/8
2	·	1/8	1/8

$p_{XY|Z=2}$

(d) Block Independent (BI):

$X \longrightarrow$, $Y \downarrow$

$Z=0$	0	1	2
0	1/8	1/8	·
1	1/8	1/8	·
2	·	·	1/2

$p_{XY|Z=0}$

$Z=1$	0	1	2
0	1/6	1/6	·
1	1/6	1/6	·
2	·	·	1/3

$p_{XY|Z=1}$

$Z=2$	0	1	2
0	1/2	·	·
1	·	1/8	1/8
2	·	1/8	1/8

$p_{XY|Z=2}$

FIG. 3. (a) is an LOPC-flagged distribution that is neither SBI nor UBI. (b) is likewise a UBI distribution that is neither SBI nor LOPC-flagged. (c) is a UBI-PD distribution that is neither UBI nor LOPC-flagged. (d) is a BI distribution that is neither UBI-PD nor UBI-PD↓. Figure 1 gives a UBI-PD↓ distribution that is not BI.

Conditional Double Markov Chain

Proposition 4 (Conditional Double Markov Chains (also Exercise 16.25 in [32]))**.** *Random variables $WXYZ$ satisfy the two Markov chains $X - YZ - W$ and $Y - XZ - W$ iff $I(XY : W|J_{XY|Z}Z) = 0$.*

Proof. If $I(XY : W|J_{XY|Z}Z) = 0$ then $I(Y : W|J_{XY|Z}Z) = 0$. The Markov chain $X - YZ - W$ follows since

$$\begin{aligned} I(XY : W|J_{XY|Z}Z) &= I(X : W|YJ_{XY|Z}Z) + I(Y : W|J_{XY|Z}Z) \\ &= I(X : W|YZ) + I(Y : W|J_{XY|Z}Z), \end{aligned}$$

where we have use the fact that $J_{XY|Z}$ is a function X and Y when given Z. A similar argument shows that $Y - XZ - W$.

On the other hand, if the two Markov chains hold, then whenever $p_{XYZ}(x, y, z) > 0$, we have

$$p(W = w|x, y, z) = p(w|x, z) = p(w|y, z). \tag{15}$$

Hence, the conditional distribution $p(w|x, y, z)$ is constant across each block $\mathcal{X}_i \times \mathcal{Y}_i$ in the maximal common partitioning of $P_{XY|Z=z}$. Consequently,

$$p_{W|XYZ} = p_{W|J_{XY|Z}Z},$$

and so for any $J_{XY|Z} = j$ and $Z = z$ for which $p(j, z) > 0$, we have

$$\begin{aligned} p(x, y, w|j, z) &= p(w|x, y, j, z)p(x, y|j, z) \\ &= p(w|x, y, z)p(x, y|j, z) = p(w|j, z)p(x, y|j, z). \end{aligned} \tag{16}$$

Thus, $I(XY : W|J_{XY|Z}Z) = 0$. □

Reversibility Conditions when $\min\{|\mathcal{X}|, |\mathcal{Y}|\} = 2$

Here we generalize Theorem 1. To do so, we will need to reference a strengthened version of Proposition 1, which first requires some new terminology. For a distribution p, let *supp*$[p]$ denote its support; the set of elements for which p assigns a nonzero probability. For two distributions p_{XY} and q_{XY} over $\mathcal{X} \times \mathcal{Y}$, we say that $q_{XY} \blacktriangleleft p_{XY}$ if, up to a permutation between X and Y, the distributions satisfy *supp*$[q_X] \subset$ *supp*$[p_X]$ and one of the three additional conditions: (i) q_{XY} is uncorrelated, (ii) *supp*$[q_Y] \subset$ *supp*$[p_Y]$, or (iii) $y \in$ *supp*$[q_Y] \setminus$ *supp*$[p_Y]$ implies that $H(X|Y = y) = 0$.

Lemma 3 ([21])**.** *Let p_{XYZ} be a distribution over $\mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ such that $p_{XY|Z=z_1} \blacktriangleleft p_{XY|Z=z_0}$ for some $z_0, z_1 \in \mathcal{Z}$. If there exists some pair $(x, y) \in$ supp$[p_{X|Z=0}] \times$ supp$[p_{Y|Z=0}]$ for which $p(x, y|z_1) > 0$ but $p(x, y|z_0) = 0$, then $K_D(p_{XYZ}) < I(X : Y|Z)$.*

Using this lemma, we are able to provide a full solution to the secrecy reversibility problem when one of the parties has a binary random variable.

Theorem 3. *Suppose that $\min\{|\mathcal{X}|, |\mathcal{Y}|\} = 2$. Then p_{XYZ} satisfies $K_D(p_{XYZ}) = K_C(p_{XYZ})$ iff p_{XYZ} is UBI-PD.*

Proof. It suffices to prove necessity. If $K_C(p_{XYZ}) = K_D(p_{XYZ})$ then by Lemma 2, $p_{XY\overline{Z}}$ must be block independent, where $I(X : Y \downarrow Z) = I(X : Y|\overline{Z})$. By the same reasoning as in Theorem 1, $K_C(p_{XYZ}) = K_D(p_{XYZ})$ implies that $K_D(p_{XY\overline{Z}}) = I(X : Y|\overline{Z})$. Hence we will apply Lemma 3 on the equality $K_D(p_{XY\overline{Z}}) = I(X : Y|\overline{Z})$ to derive a necessary condition for $p_{XY\overline{Z}}$.

Without loss of generality, assume that $|\mathcal{X}| = 2$. Since $p_{XY\overline{Z}}$ is BI, every conditional distribution $p_{XY|\overline{Z}=z}$ is either:

(I) Uncorrelated $I(X : Y|Z = z) = 0$ or

(II) Correlated and satisfying $H(X|Y, \overline{Z} = z) = 0$.

Suppose now that $H(X|Y=y)>0$ for some $y\in\mathcal{Y}$, but nevertheless y is a possible event in correlated distribution $p_{XY|\overline{Z}=z}$. The latter means that $p(x,y|z)=0$ for some $x\in\{0,1\}$. However, $H(X|Y=y)>0$ implies the existence of some $\tilde{z}\neq z$ such that $p(\overline{x},y|\tilde{z})>0$, where $\overline{x}=x\oplus 1$. With $p_{XY|\overline{Z}=z}$ having correlations, then $supp[p_{X|\overline{Z}=\tilde{z}}]\subset supp[p_{X|\overline{Z}=z}]$, and since $p_{XY|\overline{Z}=\tilde{z}}$ has either form (I) or (II), it follows that $p_{XY|\overline{Z}=\tilde{z}}\blacktriangleleft p_{XY|\overline{Z}=z}$. But then Lemma 3 implies that $K_D(p_{XY\overline{Z}})<I(X:Y|\overline{Z})$, which contradicts our assumption. Therefore, if y is a possible event in any correlated conditional distribution $p_{XY|\overline{Z}=z}$, then $H(X|Y=y)=0$. Consequently, we can define the following message for Bob and maximal conditional common functions:

$$M(y)=\begin{cases}0 & \text{if } H(X|Y=y)>0\\ 1 & \text{if } H(X|Y=y)=0\end{cases}\qquad\qquad J_{XY|\overline{Z}}(x,y,z)=\begin{cases}0 & \text{if } p_{XY|\overline{Z}=z} \text{ is uncorrelated}\\ x & p_{XY|\overline{Z}=z} \text{ is correlated}\end{cases}$$

$$J_{XY|\overline{Z}M}(x,y,z,m)=\begin{cases}0 & \text{if } m=0\\ x & \text{if } m=1.\end{cases}\tag{17}$$

It is obvious that $I(J_{XY|Z}:M|Z)=0$ since $J_{XY|Z}=0$ for all z whenever $p_{XY|Z=z}$ is uncorrelated, and $M=1$ for all z whenever $p_{XY|Z=z}$ is correlated. Also, $J_{XY|ZM}$ becomes a shared variable for Alice and Bob since it can be computed both by Alice and Bob given M. We thus, see that $p_{XY\overline{Z}}$ is UBI-PD.

□

Calculation of Theorem 2

First recall that for a two-qubit state ρ, its concurrence is defined by $C(\rho)=\sqrt{\lambda_1}-\sqrt{\lambda_2}-\sqrt{\lambda_3}-\sqrt{\lambda_4}$, where the λ_i are the non-increasing eigenvalues of the operator $\rho\tilde{\rho}$, with $\tilde{\rho}=(\sigma_2\otimes\sigma_2)\rho^*(\sigma_2\otimes\sigma_2)$ [25]. Here, ρ^* is the complex conjugate of ρ in the computational basis, and the σ_i are the Pauli matrices: $\sigma_1=\left(\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right)$, $\sigma_2=\left(\begin{smallmatrix}0&-i\\i&0\end{smallmatrix}\right)$, and $\sigma_3=\left(\begin{smallmatrix}1&0\\0&-1\end{smallmatrix}\right)$. For a two-qubit state with concurrence $C(\rho)$, its entanglement of formation is given by $\mathsf{E}(C(\rho))$, where $\mathsf{E}(x):=h(\frac{1}{2}[1-\sqrt{1-x^2}])$ and $h(x):=-x\log x-(1-x)\log(1-x)$. Note that $\mathsf{E}(x)$ is strictly convex in x.

When $|\mathcal{X}|=|\mathcal{Y}|=2$, reversible distributions are UBI-PB↓. If $K_D(p_{XYZ})>0$, then the distribution is UBI-PB, and we wish to show:

$$\begin{aligned}K_D(p_{XYZ})&=\sum_{z\in\mathcal{Z}}p(z)\mathsf{E}\left(2\sqrt{p(0|z)p(1|z)}\right)\\ E_F(\rho_{XY})&=\mathsf{E}\left(2\sum_{z\in\mathcal{Z}}p(z)\sqrt{p(0|z)p(1|z)}\right).\end{aligned}\tag{18}$$

Up to a relabeling of x, a general UBI-PB distribution in 2×2 is given by $p(x,y|z)=\delta_{xy}p(x|z)$ for $x,y\in\{0,1\}$ and arbitrary $p(x|z)$. Then $\rho_{AB}=\sum_z p(z)|\varphi_z\rangle\langle\varphi_z|$, where $|\varphi_z\rangle=\sum_{x=0}^{1}\sqrt{p(x|z)}|xx\rangle$. This corresponds to a single qubit density matrix

$$\begin{aligned}\omega&=\begin{pmatrix}\sum_z p(z)p(0|z) & \sum_z p(z)\sqrt{p(0|z)p(1|z)}\\ \sum_z p(z)\sqrt{p(0|z)p(1|z)} & \sum_z p(z)p(1|z)\end{pmatrix}\\ &=\begin{pmatrix}\sum_z p(z)p(0|z) & \frac{1}{2}\sum_z p(z)\sqrt{C(\varphi_z)}\\ \frac{1}{2}\sum_z p(z)\sqrt{C(\varphi_z)} & \sum_z p(z)p(1|z)\end{pmatrix}.\end{aligned}\tag{19}$$

It can be seen that $\sigma_2\omega^*\sigma_2=\sigma_1\omega\sigma_1$. Hence, the concurrence of ρ_{AB} can be computed from the eigenvalues of the 2×2 matrix $\omega\tilde{\omega}=\omega\sigma_1\omega\sigma_1$, which are

$$\left(\sqrt{p(0)p(1)}\pm\sum_z p(z)\sqrt{p(0|z)p(1|z)}\right)^2.$$

The Cauchy-Schwarz Inequality then gives that

$$C(\rho_{AB})=\sqrt{\lambda_{\max}}-\sqrt{\lambda_{\min}}=2\sum_z p(z)\sqrt{p(0|z)p(1|z)}=\sum_z p(z)C(\varphi_z).\tag{20}$$

Since $K_D(p_{XYZ}) = H(J_{XY|Z}|Z) = \sum_z p(z)E(C(\varphi_z))$, the calculation of Eq. (18). Note that by strict convexity of $\mathsf{E}(x)$, we have

$$\sum_{z\in\mathcal{Z}} p(z)\mathsf{E}\left(2\sqrt{p(0|z)p(1|z)}\right) = \mathsf{E}\left(2\sum_{z\in\mathcal{Z}} p(z)\sqrt{p(0|z)p(1|z)}\right)$$

iff $p(0|z)p(1|z)$ is constant for all z. This implies that $H(X|Z=z)$ is constant for all $z \in \mathcal{Z}$.